

Md Basim Uddin Ahmed

North York, ON, Canada

✉ mbahmed@yorku.ca | 📧 basimbd | 🌐 basimbd

Education

York University

M.Sc. IN COMPUTER SCIENCE

Supervisor - Dr. Song Wang

North York, ON, Canada

Sep 2024 - **Ongoing**

Shahjalal University of Science and Technology

B.Sc. (ENGG.) IN COMPUTER SCIENCE AND ENGINEERING

• CGPA - **3.73** / 4.00

Sylhet, Bangladesh

Feb 2017 - Jul 2021

Research Experience

LLM-assisted Vulnerability Detection in C/C++

MASTER'S THESIS

Ongoing

In this work, I am exploring the LLMs' capability in detecting security vulnerabilities in C/C++ projects.

- Collected **5,867 C/C++ vulnerabilities** at several abstraction levels: file-level, function-level, and statement-level.
- Explored the capabilities of 7 state-of-the-art LLMs and fine-tuned LLMs for vulnerability detection.
- Designed an LLM-agentic workflow for vulnerability detection with repo-level context.
- Publication: One benchmark paper *under-review* at **NeurIPS 2025**.
- The dataset is available at HuggingFace: <https://huggingface.co/datasets/arag0rn/SecVulEval>

Analyzing Static Analysis Tools

RESEARCH PROJECT

Ongoing

This work focuses on finding bugs in popular static analysis tools using mutation-based differential testing.

- Exploring 3 popular Java static analysis (SA) tools: ErrorProne, SonarQube, Spotbugs.
- Using LLMs to generate buggy code and their mutants which is used to test each of the SA tools.
- Found total 45 anomalies in all 3 tools. Bugs currently pending developer confirmation.
- Publication: One paper *under-review* at **ICSE 2026**.

Topic Modeling for Bengali Language

BACHELOR'S THESIS

Aug 2020 - Jul 2021

This research focused on developing a topic modeling algorithm for Bengali, and using it to extract topics from news documents. The findings of the research are a modified-LDA that works well on Bengali data, and a method to find news trends.

- Learned the LDA algorithm for topic modeling and modified it for improvement.
- Fed the Bengali dataset and tuned the hyper-parameters to fit the data.
- After the model was mature and stable, used it to analyze news trends in 9 different categories over the past 5 years.

Work Experience

Dynamic Solution Innovators Ltd.

Dhaka, Bangladesh

ASSOCIATE SOFTWARE ENGINEER

Jul 2021 - Jul 2024

Worked on the HEMS project which is a nation-wide examination management system.

- Contributed to the initial database design of the application to reflect the client's requirements.
- Implemented the result-processing module that collects, analyzes marks of around **30,000 students** per year, and prepares grades.
- Integrated 2 payment gateways that manage a total transaction amount of around **4 million BDT** per year.
- Used Java, Spring Framework, PostgreSQL as technology stack.

Synopsys Inc.

Sunnyvale, CA, USA

SOFTWARE ENGINEER (EXTERNAL)

Aug 2023 - Jul 2024

Remotely worked with Synopsys Inc. as an offshore employee. Worked on Synopsys-Detect project (now renamed to BlackDuck Detect) which is a software compositional analysis tool. It scans a software repository to find component, security, and license vulnerabilities resulting from project dependencies.

- Worked on integrating the Detect project with GitHub Actions CI/CD pipeline. Using Java, Groovy, and Gradle scripts.
- Migrated project analytics data from Google Universal Analytics to the newer Google Analytics 4 (GA4) protocol. Also, updated the code module for the new protocol.

Projects

SaliencyMix Replication

COMPUTER VISION PROJECT

Mar 2025 - Apr 2025

Replicating all the experiment results from the SaliencyMix technique proposed in ICLR 2021. The goal is to familiarize myself with Computer Vision techniques while implementing real-world experiments.

- Implemented the **ResNet** model from scratch with 3 variations (18, 50, 101). in **PyTorch**
- Implemented saliency map generation and mixing salient regions with target image.
- Used **CIFAR-10**, **CIFAR-100**, and **ImageNet** dataset.

hf_ml_apps_analysis

DATA SCIENCE PROJECT

Nov 2023

This app analyzes the ease and maintainability of Text Classification and Text Generation applications in Hugging Face by analyzing their respective project size (Lines of Code).

- Top 20 (by downloads) Text Classification and Text Generation models selected.
- All projects using those models were selected and filtered according to criteria.
- Used *Lizard* static analyzer to get the Non-comment Lines of Code (NLOC).
- Drawn conclusion by data analysis using *Pandas*, *Numpy*, *Matplotlib*, etc.

Music Genre Classification

MACHINE LEARNING PROJECT

Jun 2021

This is a machine learning based approach to predict the genre of music from an audio file.

- K-NN was used as the classification algorithm.
- Used *python_speech_features* and *Numpy* to implement the model.
- The GTZAN_dataset was used to train and test the model with a 85%–15% train–test split. The model had a 74.65% accuracy on the test set.

Skills

ML and AI Libraries

- TRANSFORMERS, PYTORCH, NUMPY, PANDAS

Programming Language & Frameworks

- PYTHON, JAVA, JAVASCRIPT & SPRINGBOOT

Database

- MYSQL, POSTGRESQL, FIREBASE

Tools & Tech

- BASH, LINUX, GIT, JENKINS, GA4

Publication

Md. Basim Uddin Ahmed, Mohammad Abdullah Al Mumin, Mahruba Sharmin Chowdhury, and Ananta Akash Podder. A systematic literature review on english and bangla topic modeling. *Journal of Computer Science*, 17(1):1–18, Jan 2021

Certifications in AI-ML

Neural Networks and Deep Learning

COURSERA

Sequence Models

COURSERA